

DOCUMENT RESUME

ED 279 212

FL 016 532

AUTHOR Chasaide, Ailbhe Ni; Davis, Eugene
TITLE A Data-Processing System for Quantitative Analysis in
Speech Production. CLCS Occasional Paper No. 17.
INSTITUTION Trinity Coll., Dublin (Ireland).
REPORT NO ISSN-0332 3889
PUB DATE 86
NOTE 32p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Acoustic Phonetics; *Articulation (Speech); Auditory
Perception; *Computational Linguistics; Computer
Software; *Data Processing; Distinctive Features
(Language); Language Research; Oral Language;
*Phonetic Analysis; Phonology; Research
Opportunities; *Statistical Analysis; Technological
Advancement

IDENTIFIERS *Trinity College (Ireland)

ABSTRACT

The data processing system used at Trinity College's Centre for Language and Communication Studies (Ireland) enables computer-automated collection and analysis of phonetic data and has many advantages for research on speech production. The system allows accurate handling of large quantities of data, eliminates many of the limitations of manual data analysis, and opens new areas of research. Possible exploitations and extensions of the current capabilities to enhance phonetic research include accurate measurement of respiratory regularities underlying stressed syllables, manipulation of signals to study aerodynamic aspects of phonation and the velocity of glottal movement, and precise synchronization of physiological, aerodynamic, and acoustic information. (MSE)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED279212

TRINITY COLLEGE DUBLIN



Centre for Language and Communication Studies

A data-processing system for quantitative analysis in speech production

Ailbhe Ní Chasaide and Eugene Davis

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

CLCS Occasional Paper No.17

ISSN 0332 3889

Autumn 1986

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

D. Little

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

2

BEST COPY AVAILABLE

FL016532

CLCS Occasional Papers

General Editor: D. G. Little

CLCS Occasional Papers report on research carried out within or in association with the Centre for Language and Communication Studies, and on research carried out elsewhere which is of special interest to the Centre's own concerns and activities. In some instances they present the texts of public lectures promoted by the Centre.

The following titles have been published and are available from The Secretary, Centre for Language and Communication Studies, Trinity College, Dublin 2.

Autumn 1981

1. D. M. Singleton, *Language transfer: a review of some recent research* (31pp.)
2. Jeffrey L. Kallen, *Linguistics and oral tradition: the structural study of the riddle* (33pp.)
3. D. M. Singleton, *Age as a factor in second language acquisition* (70pp.)

Summer 1982

4. Francis Nolan, *Voice quality and speech synthesis* (30pp.)
5. Roger Bennett, *Language elicitation procedures* (23pp.)

Winter 1982-3

6. S. M. Devitt, D. G. Little, S. P. Ó Conchúir, D. M. Singleton, *Learning Irish with Anois is Árís* (128pp.)

Autumn 1983

7. William T. Littlewood, *A communicative approach to language-teaching methodology* (19pp.)
8. Rose Maclaran, *On the interaction of semantics and pragmatics* (18pp.)
9. E. M. Harding, *Compensation strategies* (54pp.)

(continued on inside back cover)

A DATA-PROCESSING SYSTEM FOR QUANTITATIVE
ANALYSIS IN SPEECH PRODUCTION

by

Ailbhe Ní Chasaide and Eugene Davis

1 INTRODUCTION

The phonetics laboratory that has been established in the Centre for Language and Communication Studies is the first of its kind in Ireland. It thus significantly extends the scope of linguistic research that can be undertaken in this country. The possible range of instrumental/experimental phonetic research is in itself extremely broad. It embraces on the one hand the study of speech production, i.e. the physiological, aerodynamic and acoustic aspects of speech as produced by the speaker, and on the other hand the study of how speech is perceived by the listener. In this first phase of our laboratory's development we have concentrated on the area of speech production.

Much of the work in this area involves the analysis of a number of simultaneously recorded signals. These are mostly analog signals and they relate physiological, aerodynamic and acoustic parameters of speech as they vary in time. Analysis and measurement have traditionally been carried out by hand from a hard copy typically obtained on an inkwriter or ultra-violet printer. Figure 1

BEST COPY AVAILABLE

(p.5) illustrates two sample recordings, output on a Siemens Oscillomink.

This paper describes a system we have implemented in our laboratory, which permits computer-automated collection and analysis of such data. For the benefit of the non-phonetician, the following section will briefly illustrate how some of the peripheral instrumentation is used in phonetic research. The reader familiar with such instrumentation should proceed to section 3 for the description of the data-processing system.

2 PHONETIC RESEARCH APPLICATIONS

Speech is the result of complex synchronous movements of the various speech organs, set in motion by neural impulses. Lung activity provides the moving airstream (kinetic energy) which is the prerequisite of all speech. Adjustments of the vocal folds, and of the jaw, tongue and lips, provide the types of obstruction to the airstream that produce speech sounds by converting kinetic to acoustic energy. For instance, if the vocal folds are drawn loosely together in the path of the airstream, they will vibrate to give the "buzzing" sound of voicing, characteristic of vowels and of voiced consonants. Momentary closure in the mouth yields a "popping" sound (plosive) when it is suddenly released. Narrowing the air passage results in the "hissing" sound of a fricative. (The first and last sounds in the word [kaf] cash are examples of a plosive and a fricative respectively.) Once sounds have been generated they may be further modified by the size and shape of the cavities they subsequently pass through. For instance, for vowels, where the only source of sound is the voicing "buzz" produced at the larynx, the

quality we hear (e.g. whether we have an [i] or an [u] sound) will depend on the size and shape of the throat and mouth cavities, which in turn depends principally on the postures adopted by the tongue and lips.

Thus to produce even a single speech segment - say, the [t] in the word [antɪk] *antic* - fine temporal coordination of different articulatory gestures is required: raising of the jaw and tongue (providing oral closure), raising of the velum (to block off the nasal cavity, which would have been lowered for the preceding nasal sound), and opening of the vocal folds to provide voicelessness. (During the articulation of [t] the vocal folds will also begin to close again in order to be correctly positioned to produce voicing in the following vowel.) The phonetic instrumentation of interest here (some of which is illustrated below) permits one to monitor the activity and fine co-ordination of the speech organs as well as providing information on aerodynamic and acoustic characteristics of speech.

The recorded signals in Figure 1 include nasal (U_n) and oral (U_o) airflow rates, a laryngographic waveform (L_x), the speech waveform obtained from the microphone signal (Mic), the pressure beneath the vocal folds, i.e. subglottal pressure (P_s), and the waveform obtained from a photoelectric glottograph (PEG). (For an illustration of most of these inputs, see Figure 2, p.11.) To measure oral and nasal airflow rates, the subject speaks through a mask which separates airflow emanating from the nose and mouth (see Anthony 1982 for a description of this kind of equipment). The laryngographic signal is obtained from electrodes which, placed externally on the neck surface on either side of the vocal folds, measure the electrical impedance across the vocal folds (see Fourcin 1974,

Anthony 1978). The waveform reflects the very rapid vibrations of the vocal folds during voiced speech. Subglottal (and/or oral) pressure can be recorded by means of minute pressure transducers mounted on a flexible tube which is introduced through the nose (for an illustration of this equipment, see Bladon and Al-Bamerni 1982). The tip of the tube is swallowed so that the lower transducer lies in the oesophagus underneath and behind the vocal folds. (Thus strictly speaking it is measuring oesophageal pressure, which however corresponds very closely to subglottal pressure.) Above that, the other transducer is positioned in the throat, and for most articulations the pressure recorded will correspond to oral pressure. The photoelectric glottograph signal is obtained by use of a light source and photo-transistor. The light source is held externally on the neck surface underneath the vocal folds, while the photo-transistor (also mounted on a tube introduced through the nose) lies in the throat above the vocal folds. The amount of vocal fold opening determines the amount of light which penetrates to the photo-transistor. The waveform thus reflects the gross opening/closing movements of the vocal folds as, for example, with intervocalic voiceless consonants (see Andersen 1981 for an illustration of this type of data). The microphone which yields the speech waveform may be external or may be mounted inside the airflow mask, if such a mask is being used in a recording.

From these traces various kinds of measurement and deduction can be made:

- *Articulatory events* can often be ascertained either directly or indirectly, depending on the choice of signal recorded. In Figure 1, small vertically placed arrows identify a few segmentation points relating to the main ac-

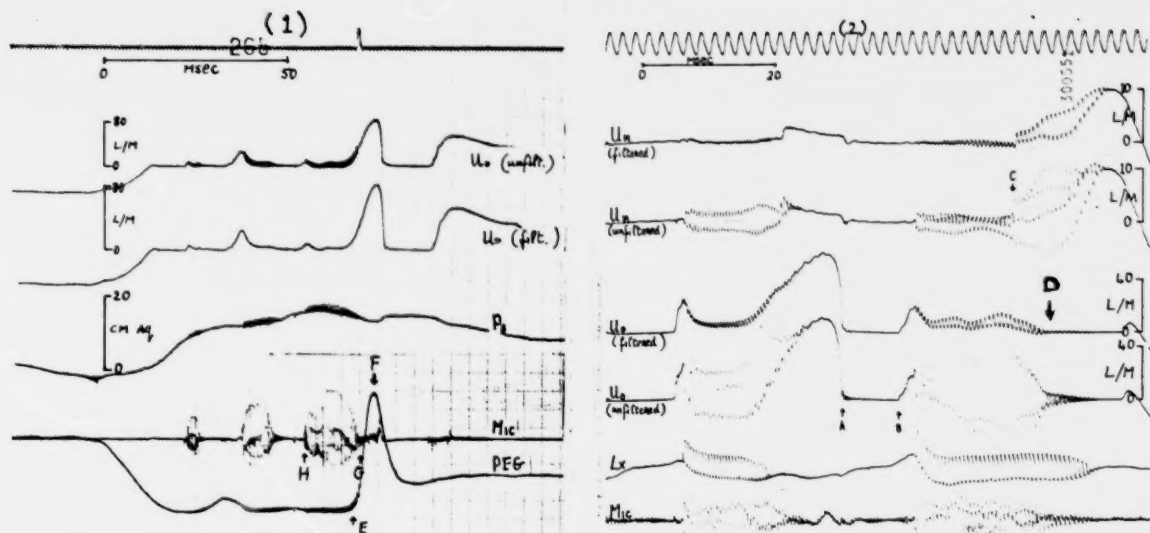


Figure 1 Inkwriter tracings showing (1) the Irish phrase "Dúirt sé brat" and (2) the last two words of the Scottish-Gaelic phrase "Thùirt e bata rium". The channels recorded and the segmentation points (↑) are discussed in the text.

tions of the vocal folds, velum and oral articulators. Points A and B identify oral closure and release respectively for the [t] in *bata* as ascertained from the drop in airflow rate to zero and the subsequent sharp rise. Point C indicates velar lowering in anticipation of the nasal segment [m], inferred in a similar manner from the nasal airflow signal. From the oral airflow trace (point D) one can tell where oral closure for the nasal occurs. The photoelectric glottograph waveform permits one to identify where the vocal folds begin to open (point E) for the voiceless plosive [t̥] in *brat*; at point F one can see that they are beginning to close again. From the speech waveform (Mic) it is possible to determine where voicing begins (point H) and ends (point G).

- Insofar as we can segment for such articulatory events in these signals, we can also make *temporal measurements*. Of interest here is not only the timing of individual articulatory events or the duration of linguistic segments, but also the fine-grained co-ordination of different articulators.
- The height of the traces indicates the *amplitudes* of the signals. Thus, for instance, direct *aerodynamic measurements* can be carried out on pressure (in units of cm H₂O) and flow (in litres/minute) signals, provided they have been calibrated for a particular experiment.

Clearly, the information and measurements which can be obtained multiply with input signals. Should one want to study intonation and stress patterns, for example, one would probably add fun-

damental frequency (F_0) and amplitude traces which are derived from the audio signal.

The type of phonetic information just described is of interest to the phonetician on two distinct if not entirely separable levels: that of describing languages and that of enhancing our basic understanding of the processes of speech production.

Linguistic description

In the first instance, the phonetician's task is to elucidate the sound structure of languages. As might be imagined, there is a wide variety of possible research avenues which might be explored, far too many to list here. Simply with the facility of making temporal measurements alone the researcher has a valuable tool to investigate suprasegmental aspects of language structure such as quantity systems in consonants and vowels, stress and rhythm. Temporal studies are also crucial to our understanding of aspects of language structure perceived as segmental, as in the case of the fine temporal adjustments that underlie phonological voicing contrasts (see Ní Chasaide 1985).

The "mix and match" possibilities of correlating very different kinds of phonetic output means that a particular aspect of language structure can, in principle, be described from a number of angles - or stages in the speech chain. Thus a study of linguistic stress might include the study of physiological and aerodynamic aspects on the one hand, durational characteristics and acoustic correlates on the other.

Inferences on speech production

Over and above the description of language structure, the phonetician seeks whatever inferences can be made regarding the processes of speech production themselves. At the peripheral level there is interest in the functioning of the speech organs - respiratory, phonatory and articulatory, and in the ways in which they interact and constrain each other's activity. Ladefoged 1967 and Anthony 1982 have investigated the regulation of the respiratory system for speech. The ways in which aerodynamic conditions and articulatory events constrain phonation have been highlighted by many investigators, such as Ladefoged 1967, Ohala 1976, and Westbury 1979. Suggestions as to how aerodynamic conditions may determine the type of glottal gesture in the control of voiceless consonants have been put forward in Ní Chasaide 1985.

More speculatively, inferences can also be made on the higher levels of speech production. Coarticulatory studies (that is, studies of the temporal overlap of the articulatory gestures for adjacent segments) have been much used as a basis for constructing and testing theories on the neural organization of speech (see, for example, Kozhevnikov and Chistovich 1965, Kent and Minifie 1977, and Fowler 1980).

One important application of this type of peripheral phonetic equipment concerns the investigation of speech pathology. This type of application may have objectives that go beyond the purely descriptive. For example, in an aerodynamic study of nasal leakage in cleft palate speech, the airflow monitoring equipment was additionally used to aid pre-operative diagnosis and post-operative assessment (cf. Anthony and Watson 1980). Similarly, studies of the intonation of deaf speech are

likely to have the additional objective of using the F₀ display as a training aid in remedial teaching (cf. McGarr et al. 1986).

Every advance in our understanding of how the speech organs function, or of the organization of speech at the peripheral and higher levels, concomitantly advances our understanding of how speech malfunctions occur and how they might be remedied. Furthermore, these insights into production constraints on speech provide many of the vital clues to the ultimate puzzle linguists seek to solve, the puzzle of why languages pattern as they do, and why they so frequently change in overwhelmingly predictable directions.

The research applications that are possible are many and varied. However, if data is all to be measured and analysed manually, this greatly restricts the quantity and sometimes the nature of the data which can in fact be handled. For example, if from every speech sample a number of complex measurements are required - say, both of amplitude and in the time domain - the cost in tedium and time dictates the quantity of data any one person can reasonably analyse. Furthermore, it can be extremely difficult to measure fine differences accurately. An example of this would be the measurement of the pitch perturbations of a few Hz which characterize the voiced/voiceless stop distinction. Other practical limiting factors can be the response frequency of the writing device such as an inkwriter when signals of a high frequency are involved. Ultra-violet (UV) recorders do handle high frequencies, but the cost of UV paper limits their likely use. A further limitation of the usual manual configuration is the difficulty of synchronizing the recorded signals with useful additional information from some device external to the system, e.g. electro-palatography or fibreoptics.

3 THE COMPUTER-AUTOMATED SYSTEM

Current configuration

Before proceeding to a more detailed description of hardware and software aspects of the system, we will first outline, with reference to Figure 2, what the general characteristics are.

Up to six channels of information can be simultaneously displayed. At present we can choose from among the following range of phonetic inputs (see Figure 1 and text of section 2):

- | | |
|-----------------------------|--|
| U_o }
U_n } | These two channels may be filtered or unfiltered. Along with the pressure monitoring equipment and amplifiers, they constitute the Gaeltec Aerodynamic Unit. |
| P_a }
P_o }
PEG } | These signals may be registered separately or together. In the latter case, we use a tube which has two pressure transducers and a photo-transistor (for the PEG) mounted appropriate distances apart. |
| F_o }
Amp } | These signals are derived from the microphone signal. |
| Mic | Either a larynx or an ordinary microphone may be used. |

The system is potentially expandable, and additional inputs could be slotted in if desired at a future time. The FM and UV recorders are currently equipped for six channels, but could be expanded to take up to 14 and 12 channels respectively.

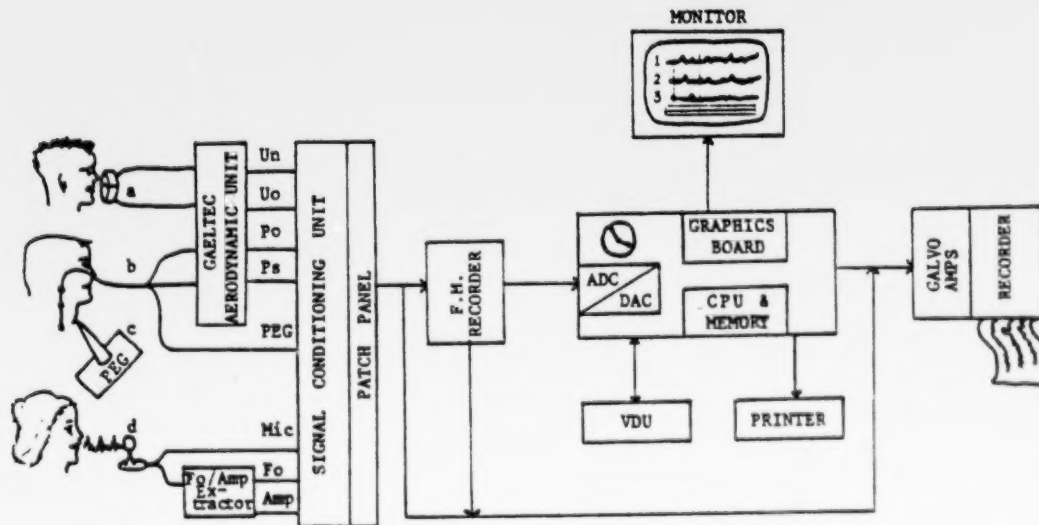


Figure 2 The general configuration of the computer-automated system. The phonetic devices illustrated are:

- airflow mask which separates oral and nasal airflow, together with the pressure transducers and amplifiers. This forms the Gaeltec aerodynamic system;
- flexible tube, on which are mounted: P_o transducer (highest); photo-transistor for PEG (mid); P_s transducer (lowest).
- light source and amplifier for PEG;
- microphone. F_o and Amplitude may further be extracted from it.

The peripheral equipment is interfaced to the computer via the signal conditioning unit. This unit fulfils the following functions:

- It houses the amplifiers which are needed to bring the analog signals to the level required for the A→D converter. It also contains overload warnings if the input signal exceeds the range of the A→D converter. Two of these channels are additionally low-pass filtered to 1000Hz. This prevents aliasing in cases where signals which have high frequencies (such as the microphone signal) are sampled at comparatively low sampling rates.
- The unit also allows the baseline of any signal to be offset. Such an adjustment is necessary in order to ensure that the full range of the A→D converter is used. The range is from +2.56V to -2.54V, and would otherwise be underused in the case of signals which only have a positive voltage output (or, indeed, for the airflow channels, where ingressive flow which corresponds to the negative range might not be of interest in many applications).
- The signal processing unit interfaces each individual item of equipment with the rest of the system. It has a 52-way patch-panel with jack plugs to which all inputs and outputs are connected. When all patch-cords are removed, we get a standard configuration of phonetic equipment, which we thought would be the most popular and frequently used. Using patch-cords the signal can be distributed in any configuration. This particular facility is very useful for teaching, as

it makes for speed and flexibility in switching between configurations.

- The signal conditioning unit also houses the electronics that enable the computer to control the FM and UV recorders. Data collection is computer controlled, and all instructions are given via the keyboard.

Analog to digital conversion can be carried out on 1, 3 or 6 channels. In addition to the number and choice of these channels, the user chooses from a range of possible sampling rates (which determine the length of any one speech sample that can be input).

The digitized signals are displayed on a monitor. For stability and persistence of display, the monitor uses a green phosphor tube.

Two types of data display are possible, the choice depending on the purpose for which it is intended:

- One can obtain a display in real time, rather like an oscilloscope. This facility is useful for teaching/demonstration purposes, as well as for setting up equipment. The screen display can be held at any point and saved to hard disk if so desired. In a teaching situation this allows one to hold in readiness illustrations of particular types of articulations. Measurements cannot, however, be carried out on this type of display.
- The other type of display is designed for research purposes. It is not carried out in real time, but rather in chunks whose size depends on the number of channels and the sampling rates chosen. The time reference

as well as the calibration coefficients can be accessed for this display-type, and therefore measurements of signal amplitudes and of durations can be made.

Calibration of the signal conditioning unit takes into account what the zero offset of each channel might be. Signals such as airflow/pressure are measured manually with manometers, and the calibration coefficients for these individual channels are entered via the keyboard.

The screen display can be manipulated in a number of ways:

- Measurement of signal amplitudes and of durations is carried out with the use of two vertical cursors which can be moved across the screen. The numerical results are shown in the lower part of the display.
- The number of channels to be displayed can be changed from 6 to 3, or to 1.
- The sequence of these channels (i.e. the order in which they appear on screen) can be changed at will.
- The display can be reduced or expanded. The maximum expansion possible will depend ultimately on the sampling rate of the digitized channel(s).

The UV recorder can provide a hard copy of the display either during or after an experiment.

Hardware

The system is configured around a Cromemco System III microcomputer with a Z80 CPU and hard disk of 10Mb capacity. Advances in 16-bit micro-technology mean that the Z80 chip is now outdated. However, the choice of the Cromemco was prompted by the fact that it offered the best microcomputer graphics potential in terms of resolution and speed. As the graphics board contains its own 96K of RAM it does not slow down the main processor, which would happen if graphics memory were part of the user memory.

Our Cromemco system was configured in the following way (see illustration in Figures 3a and 3b). There is 256K of RAM arranged into 4 x 64K banks, one of which is taken up with the operating system (this is CROMIX, a UNIX lookalike). The program modules relevant to the collection, display and analysis of data can be called up into the other three switchable databanks of 64K, and this allows for branching between program modules. For example, if when setting up the equipment parameters one wished to go back and recalibrate, one could "hold" the current module while jumping to the calibration module. The graphics RAM is divided into two blocks of 48K each. For graphics, one of these 48K RAM boards is used as graphics memory; the other one is not required for our application, so we use it to store the data from the A→D conversion. The configuration also includes a data conversion board which has 7 channels A→D as well as 7 channels D→A and a parallel input/output (I/O) port. (The parallel I/O port is connected to the signal conditioning unit (SCU) which controls the FM and UV recorders.) There is also a frequency clock module used for time reference, which enables us to know where we are (time-wise) in the digitized data.

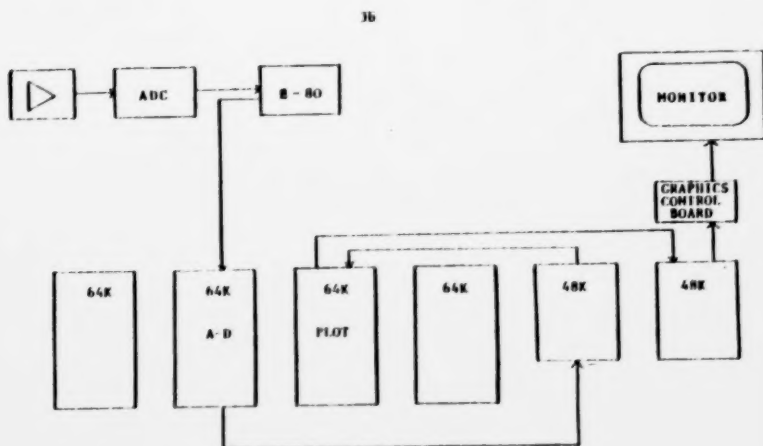
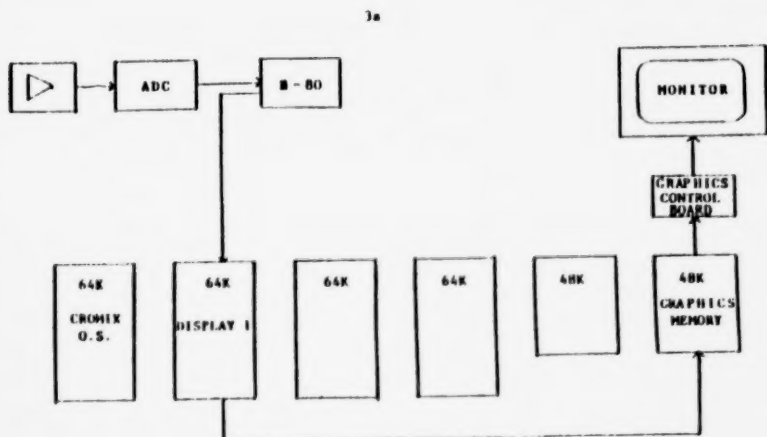


Figure 3a & 3b: The use of RAM for the two types of graphics display
(see also Figure 4)

3a program module for real time display (DISPLAY 1)

3b program module for display and analysis (DISPLAY 2)

Software

The software is written in a combination of assembly language and Pascal, and is designed so that different aspects of data collection, display and analysis are handled in separate program modules. These modules fit together neatly, yet remain sufficiently decoupled to allow them to be independently modified. The program module flow is as illustrated in Figure 4.

The initial program module, labelled PROCESS, simply allows the user to select one of the other modules. The CONTROL module gives the user keyboard control of F.M. and U.V. recorders via the SCU, and allows one to select the phonetic inputs for an experiment if these differ from the standard configuration. The CALIBRATION module allows the user to calibrate the SCU for each channel. As was mentioned earlier, the SCU permits the baseline of any channel to be offset, so that maximal use can be made of the A→D converter's range. The calibration program takes this into account and stores on file information relating to the zero and gain of each channel. Furthermore, in this program module the particular calibration coefficients for individual channels are entered, including the units of calibration (e.g. litres/minute). (These coefficients are obtained for each experiment by using manometers etc.) Finally, in this program each channel is labelled if the configuration used differs from the standard configuration.

There are two modes of operation for the graphics display. The first, DISPLAY 1, is a real-time display of the analog signals. It functions much like an oscilloscope, and is especially intended for teaching and demonstration purposes. In this mode, the converted signals are read straight through to the graphics memory (see Figure 3a),

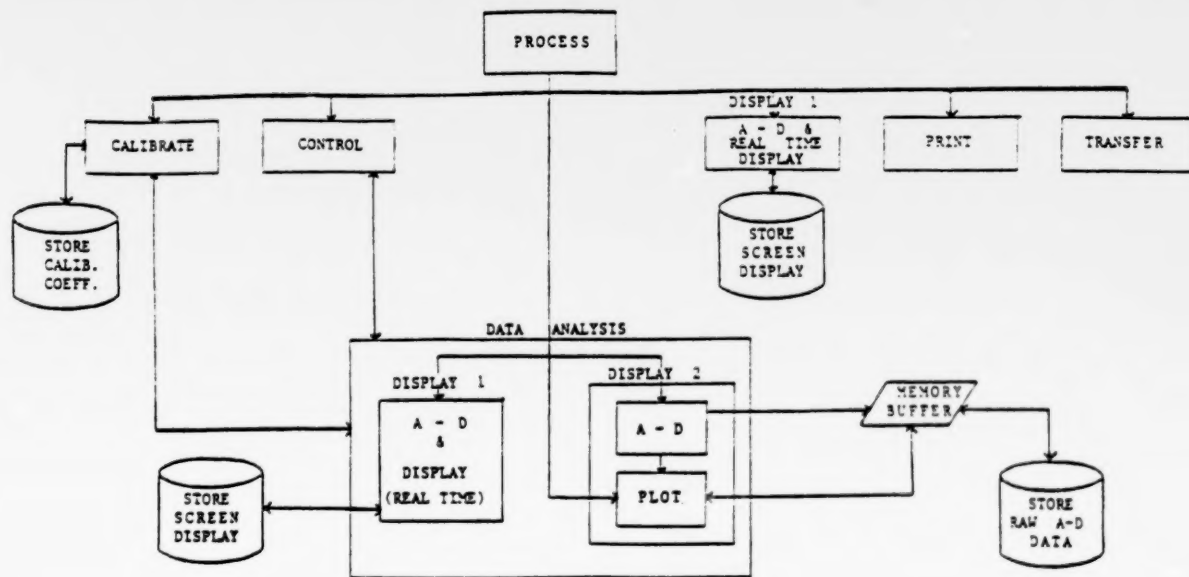


Figure 4 Program module flow

 = hard disk
 = program module

and the traces scroll across the screen in an instantaneous fashion. The screen display can be held at any particular moment, and if so desired, stored on disk. The data of this type of "screen-shot" cannot be manipulated or measured. However, it can be useful for teaching to have a number of such stored "shots" illustrating particular articulations, e.g. airflow traces to show the difference between plosives and implosives or between, say, cleft palate and normal stop articulations.

The second graphics display mode, DISPLAY 2, forms part of the larger program module, DATA ANALYSIS, used for the measurement and analysis of data for research. In DISPLAY 2, the data are converted and then stored in the spare 48K block of screen memory (see Figure 3b). Hence it can be manipulated and measurements can be obtained. Using the calibration information the amplitude of each value can be calculated. Likewise, information from the frequency clock makes it possible to calculate the time dimension. In this mode of operation the data would typically be recorded initially on FM tape and appropriate sections afterwards selected and analysed. The digitized data held in the 48K block of memory can be stored on hard disk, and can be used for further analysis at a later stage.

Nested within the larger program module DATA ANALYSIS we have a further sub-module, DISPLAY 1, identical to the similarly labelled but separate program module mentioned above. Its inclusion here is specifically intended to facilitate the setting up of signals prior to analysis in DISPLAY 2.

There are two remaining program modules, PRINT and TRANSFER. The former simply allows one to print out the results of data collected, while

the latter allows a link to the College's mainframe computer for the transfer of raw data, should one wish to use the statistical packages on the mainframe.

Resolution

The A→D converters have a resolution of 8 bits (0-255 levels). This effectively means that when the full range of the converter is used the resolution accuracy is 0.39%. This was considered more than adequate for the present application.

The user selects the desired sampling frequency for the conversion by choosing an appropriate ADC Delay Timebase setting from a range of 1 (highest rate) to 99 (lowest rate). As illustrated in Table 1, the choice of the ADC Delay Timebase setting in conjunction with the number of channels to be converted determines the actual sampling rate for the individual channels. It also determines the length of speech sample which can be processed in one go (in other words, which can be held in the 48K storage RAM). The maximum expansion that is possible in the display of data is a function of the recording sampling rates and the graphics resolution and will be explained below.

The graphics resolution is 754 horizontal and 482 vertical dots. The screen is divided into two areas in the vertical dimension:

5/7ths - reserved for displaying the signals

2/7ths - reserved for displaying text and numerical results

The user can choose to display 1, 3 or 6 channels, and the number of dots available in the vertical

TABLE 1 The relationship between sampling rates, number of channels, and the length of speech sample for 3 of the possible 99 ADC settings. The rightmost column indicates the maximum expansion possible in each case.

ADC Delay Timebase	No. of channels	Sampling rate (per channel) KHz	Sample length ms	Max expansion (1 screenful=) ms
1	1	33.37	1,465.40	21.82
1	3	16.97	960.49	42.00
1	6	9.77	834.33	74.46
50	1	5.30	9,230.10	137.45
50	3	4.59	3,548.93	158.40
50	6	3.83	2,128.55	189.96
99	1	2.87	17,019.50	253.45
99	3	2.65	6,145.50	274.29
99	6	2.38	3,426.84	305.83

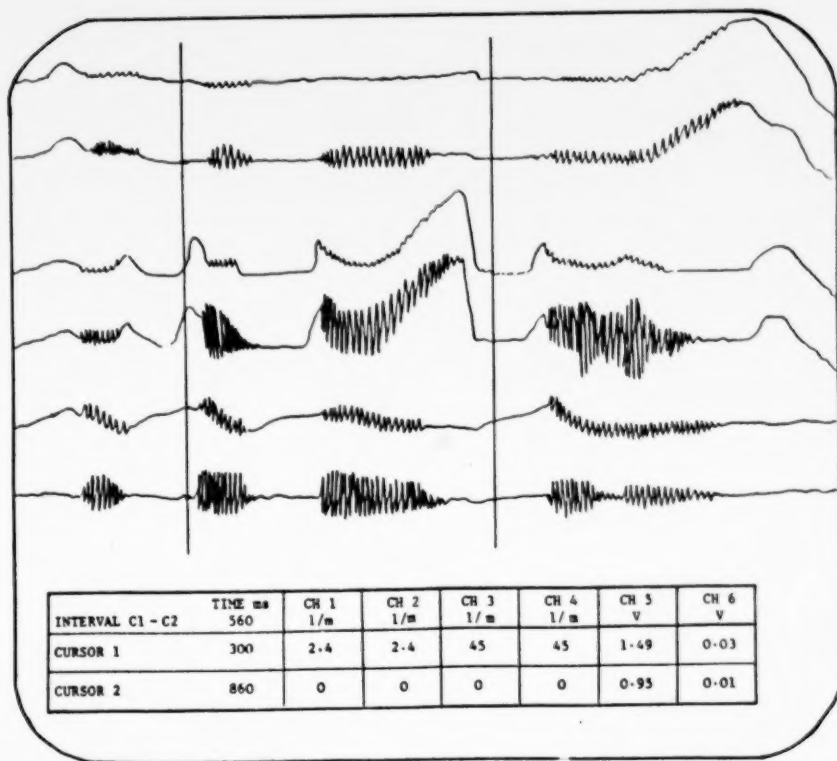
dimension for each channel will vary accordingly. When all six channels are in operation one has 70 dots per channel.

Manipulation of display

Figure 5 gives a schematic representation of a screenful of data. As can be seen, the signals with labels are displayed in the upper part of the screen, while numerical results are shown in the lower part. The screen display has the following features:

- It can be expanded and contracted. Expansion can proceed either from the beginning of the displayed data or from the location of either of two cursors (see below). Thus, for example, if we wish to look in greater detail at a particular section of the display, we can move a cursor up to that point and replot from there using an expanded timebase.
- In a fashion similar to the selection of a delay timebase at the recording stage, the user selects a display timebase from a range of 1 (= maximum expansion) to 99 (= maximum compression). When a timebase of 99 is chosen, 99 data points are plotted per vertical line, whereas with the timebase 1 a single data point is plotted per vertical line. The overall maximum expansion possibilities are therefore dependent ultimately on the sampling rates of the recording, and thus on the number of channels and the delay timebase chosen at the recording stage. The rightmost column in Table 1 tabulates the maximum expansion that can be obtained for a range of recording settings.

Figure 5 Schematic representation of one screenful of data



- Two cursors can be moved across the screen. Using the < and > keys, the cursors move slowly across the screen from one sampled data point to the next. To obtain rapid cursor movement the arrow keys are used; the greater speed is due to the fact that now the cursors are moving between vertical display points.
- In the lower part of the screen the different channels are listed. The amplitudes of each channel at the points where cursor 1 and cursor 2 are placed are listed beneath boxes detailing the calibration units used. The time location of each cursor, as well as the time difference between the two, are also listed. These values are changed automatically once a cursor is re-positioned.
- The user can select the number of channels to be displayed and the order in which the channels appear on the screen. Thus, for example, the positions of channels 6 and 2 in this figure might be reversed. This facility can be useful when dealing with fine details in segmentation.

4 CONCLUSIONS

The system described, enabling computer-automated collection and analysis of phonetic data, confers many advantages on the researcher of speech production. By allowing much greater quantities of data to be handled more accurately, it eliminates many limitations imposed by a manual approach, limitations which effectively restrict the research which might be possible in principle (some of which is illustrated in section 2). In addition, computer-automation of the process opens up new vistas of research which would not otherwise be possible. To conclude, we will mention some ways in which the scope of phonetic research would be enhanced by exploiting and building on our current system; these are specific avenues which we are now hoping to explore in our laboratory.

First of all, by virtue of the fact that the signals are stored in digitized form, they can be manipulated in ways that would be difficult or impossible if one were restricted to a hard copy from an inkwriter or UV printer. For example, it would be a simple task to derive from the U_o signal (denoting rate of oral air flow) a further measurement corresponding to the volume of air expended over a particular stretch of speech, the beginning and end of which could be defined by the location of the cursors. This is something we would like to do as a means of investigating the possible respiratory regularities that may underlie stressed syllables. Also, to study aerodynamic aspects of phonation, it would be possible to derive from oral and subglottal pressure signals a third signal corresponding to the difference between them: this is the transglottal pressure drop, a vital factor in the regulation of phonation. Numerous possibilities exist for such useful manipulation of signals. To mention just one more, it is possible to derive from

the waveform of the photoelectric glottograph (which shows vocal fold movement) a signal relating to the velocity of glottal movement (see, for example, Löfqvist and Yoshioka 1981).

In addition to the manipulation of stored signals to derive further information, the present system would make it relatively easy to synchronize our analog signals with information from external equipment such as, for example, a high-speed cine camera. This should enable one to correlate, say, the frames of a film showing vocal fold movement with our recordings of physiological, aerodynamic and acoustic information.

The Cromemco system can also be expanded to use another processor with its own memory, running an independent program. This could be used, for example, to collect electropalatographic data, sampling each time an A→D conversion is done. The analog data converted by the main processor could thus be synchronized with the electropalatographic data. The display, showing the palate and areas of articulatory contact, would appear along with the analog signals (either at the top left-hand corner of the screen or, alternatively, on a separate monitor), and would update automatically as a cursor was moved across the screen. This would allow us to correlate our analog signals with precise articulatory information which might not otherwise be inferrable. This particular development is one we hope to implement in the near future.

REFERENCES

- Andersen, P., 1981: "The effect of increased speaking rate and intended loudness on the glottal behaviour in stop consonant production as exemplified in Danish p", *Annual Report of the Institute of Phonetics, University of Copenhagen* 15, pp.103-47.
- Anthony, J. K. F., 1978: "The Volscope", *Health Services Bulletin* (Scottish Home and Health Department, Edinburgh) 36.3, pp.321-8.
- Anthony, J. K. F., 1982: "Breathing and Speaking. The Modification of Respiration for Speech." Ph.D. thesis, University of Edinburgh.
- Anthony, J. K. F. and A. C. H. Watson (1980): *Mid-project Report on the Clinical and Phonetic Investigation of Cleft-palate Speech*. Edinburgh: Scottish Home and Health Department.
- Bladon, R. A. W. and A. Al-Bamerni, 1982: "One-stage and two-stage temporal patterns of velar coarticulation", *Journal of the Acoustical Society of America* 72, S104.
- Fourcin, A., 1974: "Laryngographic examination of vocal fold vibration". In Wyke (ed.), *Ventilatory and Phonatory Control Systems*. Oxford University Press.
- Fowler, C., 1980: "Coarticulation and theories of extrinsic timing", *Journal of Phonetics* 8, pp.113-33.
- Kent, R. D. and F. D. Minifie, 1977: "Coarticulation in recent speech production models", *Journal of Phonetics* 5, pp.115-35.

- Kozhevnikov, V. A. and L. A. Chistovich, 1965: *Rech Artikulyatsya: 1. Vospriyatie*. Moscow and Leningrad. Translated as *Speech: Articulation and Perception*. Springfield, Va.: Joint Publications Research Service, United States Department of Commerce.
- Ladefoged, P., 1967: *Three Areas of Experimental Phonetics*. Oxford University Press.
- Löfqvist, A. and H. Yoshloka, 1981: "Laryngeal activity in Icelandic obstruct production", *Nordic Journal of Linguistics* 4, pp.1-18.
- McGarr, N., M. A. Head, M. Friedman, A. M. Behrman and K. Yondelman, 1986: "The use of visual and tactile sensory aids in speech production training: a preliminary report", *Journal of Rehabilitation Research and Development* 23.1, pp.101-9.
- Ní Chasnaide, A., 1985: "Preaspiration in Phonological Stop Contrasts". Ph.D. thesis, University College of North Wales, Bangor.
- Ohala, J. J., 1976: "A model of speech aerodynamics", *Report of the Phonology Laboratory, Berkeley* 1, pp.93-107.
- Westbury, J. R., 1979: *Aspects of Temporal Control of Voicing in Consonant Clusters in English*. *Texas Linguistic Forum* 14. Austin: University of Texas, Department of Linguistics.

© 1986
University of Dublin
Trinity College

Autumn 1984

10. Jeffrey L. Kallen, *Generative phonology in the clinic* (38pp.)
11. Gearóid Ó Ciaráin, *The affective dimension in second/foreign language learning: an interactional perspective* (42pp.)

Spring 1985

12. D. M. Singleton and D. G. Little, *Foreign languages at second level: defining a syllabus, with particular reference to the needs of the senior cycle* (19pp.)
13. John Harris, *The polylectal grammar stops here* (12pp.)

Spring 1986

14. D. G. Little and A. J. Grant, *Learning German without a teacher. Report on a self-instructional programme for undergraduate students of Engineering Science at Trinity College, Dublin, 1982-84* (60pp.)
15. Máire Owens, *Eithne: a study of second language development* (62pp.)
16. D. Wilson and D. Sperber, *Pragmatics: an overview* (36pp.)

Autumn 1986

17. Ailbhe Ní Chasaide and Eugene Davis, *A data-processing system for quantitative analysis in speech production* (28pp.)
18. Seán M. Devitt, *Learning a foreign language through the media* (69pp.)
19. Meriel Bloor and Thomas Bloor, *Languages for specific purposes: practice and theory* (34pp.)